OSA Imaging and Applied Optics Congress 2021 (3D, COSI, DH, ISA, pcAOP) © OSA 2021

End-to-End Optimized Adversarial Deep Compressed Super-Resolution Imaging via Pattern Scanning

Kangning Zhang[1*], Junze Zhu[2*], Weijian Yang[1†]

[1] *Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA*

[2] *Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, Anhui, 230026, P.R.China*

[*] *Equal contribution,* [†] *wejyang@ucdavis.edu*

Abstract: We propose an end-to-end optimized adversarial deep compressed imaging modality. This method exploits the adversarial duality of the sensing basis and sparse representation basis in compressed sensing framework and shows solid super-resolution results.

1. Introduction

The growing demand of high-speed imaging in various areas such as biomedicine, remote sensing and consumer electronics has called for new development of imaging modality and reconstruction algorithms. Imaging through compressed sensing (CS) [1-3] has drawn substantial attention due to the rapid progress on computational algorithms and deep learning techniques. We recently developed a new imaging modality termed deep compressed imaging via optimized pattern scanning (DeCIOPS) [4], where we projected and scanned an optimized illumination pattern on the object and collected the sampling signal using a single-pixel detector. We then reconstructed the object using a CS-inspired neural network. Using an end-to-end optimization framework, we jointly optimized the illumination pattern and the reconstruction network. Thanks to the fast scanning speed and the end-to-end optimization, this new imaging approach significantly increases the imaging speed than the typical switching-mask based single-pixel camera, while retaining a high reconstruction quality. In this paper, we propose a new reconstruction algorithm based on generative adversarial network, which further enhances the reconstruction quality of DeCIOPS. Our new algorithm optimizes the pair of the sensing basis (related to the illumination pattern) and sparse representation basis (related to image reconstruction) by exploiting the adversarial duality between them. Guided by the compressed sensing theorem, we introduce the general Jensen-Shannon Divergence (JSD) into the loss function to maximize the incoherence of the pair. This strategy enhances the reconstruction quality. We term this as adversarial deep compressed imaging (ADCI). Like DeCIOPS, ADCI enables an end-to-end optimization of the illumination pattern and reconstruction algorithm. ADCI shows reliable super-resolution results and outperforms other state-of-the-art single-image super-resolution (SISR) methods such as DCSRN [5], SRGAN [6] and ISTA-Net+ [7].

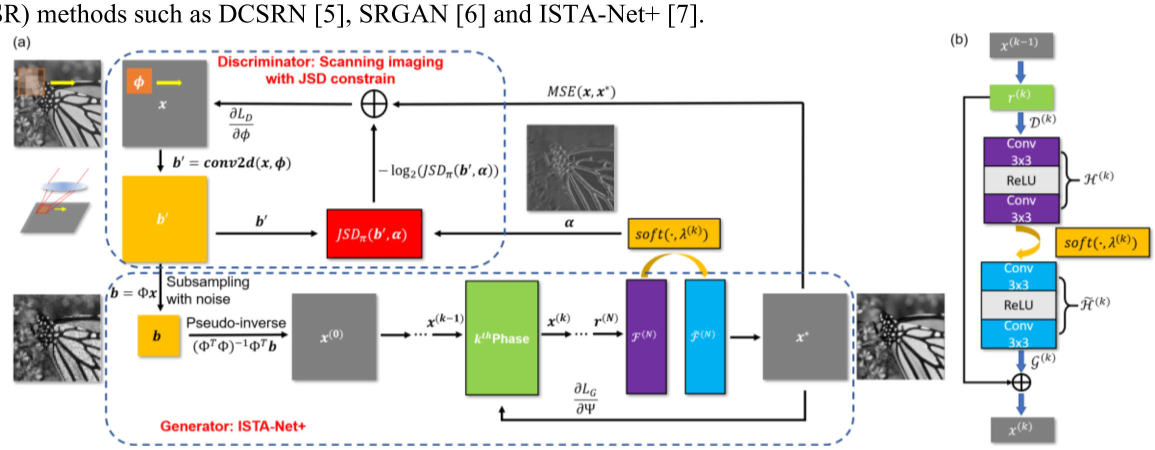

Fig. 1. (a) Schematic of the end-to-end optimization of the illumination pattern and the adversarial framework. (b) Architecture of the k^{th} phase in ISTA-Net+, where $\boldsymbol{r}^{(k)} = \boldsymbol{x}^{(k-1)} - \rho\Phi^T\left(\Phi\boldsymbol{x}^{(k-1)} - \boldsymbol{b}\right)$, $\mathcal{F}^{(k)} = \mathcal{D}^{(k)} \circ \mathcal{H}^{(k)}$ and $\tilde{\mathcal{F}}^{(k)} = \tilde{\mathcal{H}}^{(k)} \circ \mathcal{G}^{(k)}$. Ψ represents all the trainable parameters in the generator.

2. Principle of the image formation and object reconstruction

We simultaneously update the illumination pattern and a CS-inspired deep neural network under the framework of an adversarial end-to-end model (Fig. 1a). The pattern scanning encodes the high-resolution object $\boldsymbol{x}$ into a low-resolution measurement $\boldsymbol{b}$ by a subsampled 2D-convolution transfer matrix Φ. Then, an N-phase ISTA-Net+ generates a reconstruction $\boldsymbol{x}^*$ from the measurement. Each phase (Fig. 1b) contains an architecture-symmetric pair of a forward transform $\mathcal{H}^{(k)}$, a backward transform $\tilde{\mathcal{H}}^{(k)}$, linear operators $(\mathcal{D}^{(k)}, \mathcal{G}^{(k)})$, and a soft shrinkage threshold motivated by the conventional ISTA (Iterative Shrinkage-Thresholding Algorithm). The loss function of the generator L_G (Eq. 1) calculates the mean-square-error (MSE) between the reconstruction results and the ground truth with a constraint of $\tilde{\mathcal{H}} \circ \mathcal{H} = I$ weighted by γ. The loss function of the discriminator also depends on the MSE but with a

JSD constraint (Eq. 2). In the JSD measure, π and $1-\pi$ are the weights of the two normalized signal distributions. The two signals b' and α are the representation of sensing basis Φ and the sparse representation respectively. We trained the model with 500 samples from ImageNet [8] and test the model with 79 samples from two widely used benchmark datasets: Set11 and BSD68 [7] on GPU RTX2080Ti 11GB.

$$L_G = L_{error} + \gamma L_{constraint} = \|\boldsymbol{x}^* - \boldsymbol{x}\|_2^2 + \gamma \left(\sum_{k=1}^{N} \left\| \widetilde{\mathcal{H}}^{(k)} \left(\mathcal{H}^{(k)}(\boldsymbol{x}) \right) - \boldsymbol{x} \right\|_2^2 \right) \quad (1)$$

$$L_D = -\log_2(JSD_\pi) + \lambda L_{error}, \text{ with } JSD_\pi = \pi \sum_i b_i' \log_2 \left(\frac{b_i'}{\pi b_i' + (1-\pi)\alpha_i} \right) + (1-\pi) \sum_i \alpha_i \log_2 \left(\frac{\alpha_i}{\pi b_i' + (1-\pi)\alpha_i} \right) \quad (2)$$

3. Simulation results

We compare our proposed ADCI with two conventional non-learnable algorithms and five state-of-the-art single-image super-resolution deep learning methods. All methods were trained in the end-to-end framework to jointly optimize the illumination pattern and the corresponding reconstruction algorithm (excluding the two non-learnable algorithms). For a fair comparison, each method was trained independently with the same number of epochs. ADCI is able to reconstruct more details of the image with sharper edges (Fig. 2), with a higher PSNR/SSIM for different subsampling ratio (Table 1). This illustrates the effectiveness of the adversarial induced compressed imaging modality.

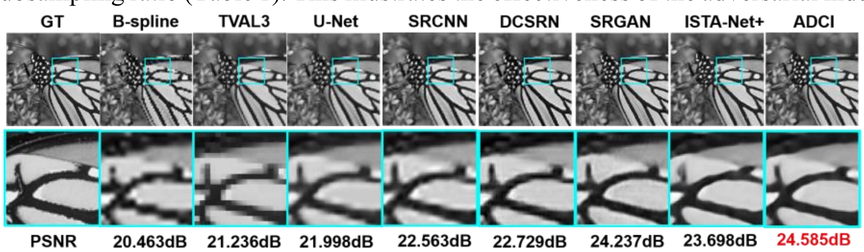

Fig. 2. Simulation results of a representative sample *butterfly* by an end-to-end trained model with reconstruction algorithms of B-spline, TVAL3, U-Net, SRCNN, DCSRN, SRGAN, ISTA-Net+, and ADCI at a subsampling rate of 6.25%. GT, ground truth.

Table.1. Average PSNR (dB)/SSIM of the reconstruction results among different end-to-end trained super-resolution algorithms and ADCI

	B-spline	TVAL3	U-Net	SRCNN	DCSRN	SRGAN	ISTA-Net+	ADCI
25% Subsampled	21.26/0.54	23.42/0.65	24.30/0.73	25.68/0.81	24.89/0.75	26.91/0.83	26.57/0.83	**27.19/0.84**
11.11% Subsampled	20.84/0.50	21.92/0.56	22.66/0.63	22.92/0.65	23.37/0.67	23.87/0.69	23.80/0.71	**24.67/0.73**
6.25% Subsampled	19.82/0.48	21.33/0.53	22.23/0.58	22.41/0.61	22.70/0.62	23.11/0.64	22.74/0.64	**23.40/0.66**
4% Subsampled	19.64/0.43	20.54/0.51	21.03/0.52	21.07/0.51	21.49/0.54	22.12/0.57	22.04/0.59	**22.47/0.60**
Trainable parameters	None	None	371477	345565	288017	362040	337010	337010

We explore the principle of the adversarial model. In compressed sensing theory, a greater incoherence between sensing basis and sparse representation basis means a more efficient data acquisition [1]. Incoherence describes how unlike the pair of signals are, so we use JSD as a constraint of loss in the discriminator. By a careful design of the loss function (Eq. 2), we converge to a higher value of JSD and thus a larger incoherence. In Table 2, ~58.3% greater JSDs are generated by ADCI compared with a pure ISTA-Net+, explaining why ADCI outperforms ISTA-Net+ (Table 1).

Table.2. Comparison of JSD between end-to-end trained model of the pure ISTA-Net+ and ADCI

	ISTA-Net+	ADCI
25% Subsampled	0.0460	**0.0630**
11.11% Subsampled	0.0502	**0.0793**
6.25% Subsampled	0.0427	**0.0753**
4% Subsampled	0.0461	**0.0755**

In summary, we propose an end-to-end optimized adversarial deep compressed imaging model which can achieve a highly efficient data acquisition and object restoration. ADCI outperforms other algorithms in SISR task and holds great promise in high-speed imaging.

Acknowledgement We acknowledge support from National Science Foundation (CARRER 1847141) and Burroughs Wellcome Fund (Career Award at the Scientific Interface).

References

[1] E. J. Candès, and M. B. Wakin, "An introduction to compressive sampling," IEEE signal processing magazine 25, 21-30 (2008).
[2] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," IEEE signal processing magazine 25, 83-91 (2008).
[3] G. M. Gibson, S. D. Johnson, and M. J. Padgett, "Single-pixel imaging 12 years on: a review," Opt Express 28, 28190-28208 (2020).
[4] Kangning Zhang, Junjie Hu, and Weijian Yang, "Deep compressed imaging via optimized pattern scanning," Photon. Res. 9, B57-B70 (2021).
[5] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), pp. 4700-4708.
[6] J. J. D. Khoo, K. H. Lim, and J. T. S. Phang, "A Review on Deep Learning Super Resolution Techniques," in *2020 IEEE 8th Conference on Systems, Process and Control (ICSPC)* (2020), pp. 134-139.
[7] J. Zhang, and B. Ghanem, "ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing," in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 1828-1837.
[8] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition* (2009), pp. 248-255.